WebGIS Application for House Discovery

1st Shengyi Zhang
1097063
Lakehead University
zhangs@lakeheadu.ca

2nd Beili Yin
1148875
Lakehead University
byin@lakeheadu.ca

***Abstract*—For this project, we developed a program based on WebGIS for House Discovery. Reviewed the history of Geographic Information System, explained and completed the core functions of the system, database, interface, and interface design so that users can safely log into our system, query the house information of their choosing.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

With the rapid development of the Internet and mobile applications, GIS has gone from desktop map editing software operated by professionals to WebGIS applications that can be seen everywhere today [2]. For example, map navigation software has become an indispensable application in our daily lives. Usually what we view and connect is discrete data, which does not contain location information. The data processed by Geographic Information System is the data that generally contains location information such as latitude and longitude. GIS data formats can be divided into vector data and raster data [3]. Vector data is Generated by points, lines and surfaces, can be enlarged without loss, while raster data is usually composed of common sliced pictures such as png format.

This project aims to develop an application for finding a house information through a map, combining relational data and GIS data to provide users with a multi-angle house selection plan. The report mainly solves and realizes the analysis of requirements and the module design according to the requirements, including how to divide the system into several modules, explaining the interface between each module, the messages transmitted between modules, and the design of database structure. This report will give a detailed description of the design of this system.

Identify applicable funding agency here. If none, delete this.

II. Problem Review

Traditionally, people use paper media and intermediaries to publish and view housing information, which is relatively inefficient [4]. Nowadays, people generally search and view housing through websites and online maps. For example, the two largest online house information websites in China and the United States, such as trulia and beike. Real estate trading platform, but after several trials and researches, it is found that such platforms cannot meet the requirement of finding houses by viewing nearby facilities and buildings, so we decided to develop a bus, hospital, school, shopping mall, office buildings, etc. within a certain distance through the annex to filter out housing information retrieval applications. The overall process is to use spring to develop a crawler that regularly crawls real estate website data every day. Here, we take the listings in Lianjia.com-Shanghai area as an example.

The data is scrapped from Lianjia and then stored in the database. In addition, we used Java Springboot, hibernate, Spring Security for software development . The permission controller end is used to process front-end requests and return restful API. The front end selects the LayUI [1] framework for web page development. The front-end and back-end of our project were developed in separated Model-View-Controller-Service architecture.

This application can count and display key information such as the number of houses available for sale in all districts in Shanghai, the average price per square meter, the average total price and other key information. Then the user can click on any district, select the query distance, and search conditions, on the map Mark the buses, hospitals, schools, shopping malls, and office buildings attached to the community, and update the query results to the database simultaneously.

Based on the collected information and the different data tables established, the house analysis function is realized, and the radar chart is used to present it.

Based on the collected housing information with geographic information, the distribution of housing prices in different areas is displayed on the map in the form of a heat map.

III. Research Methodology

After detailed research and survey, there are three different types of structures as below (in Fig 1).We used open-source software to build our system.

A. *Business intelligence strategy*

- Target:

  We want to build an application for people who want a house can check and compare the house price of different distinct in a single map as below (in Fig 2) [5].
- Advantages:

  1. Search house directly from the map, and people can add different conditions to filter their target's house, and see all the houses' price in a single map.

  2. The surrounding buildings, facilities such as schools, subways, restaurants and transports can be easily viewed and compared by zoom in and zoom out.
- Competitor analysis:

  1. Trulia [6]: In America, there is a popular app named trulia, we downloaded and tested this app, and found it has similar functionalities but doesn't have search by map function and it separates the map view, and schools, shop and eats, etc into the different map. It supports many modules, such as tools for people to calculate the mortgage, display the price history, and recommend similar homes(Recommendation system), summarize the tags and reviews of the house(NLP).

  2. Beike: In China, there is a popular app named Beike, we used it before and found it very convenient and support search houses by map. We want to build a similar application for the Canadian real estate market, the main

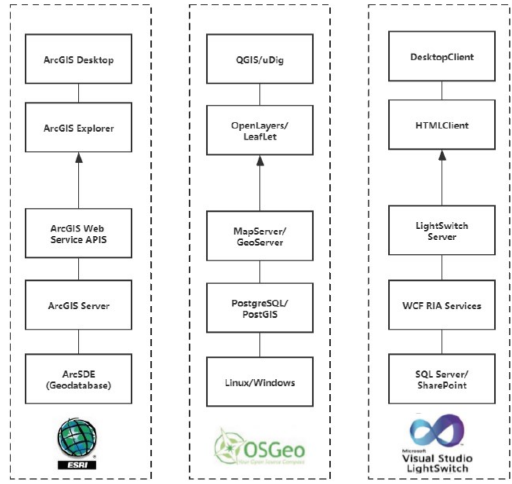

Fig. 1. Survey

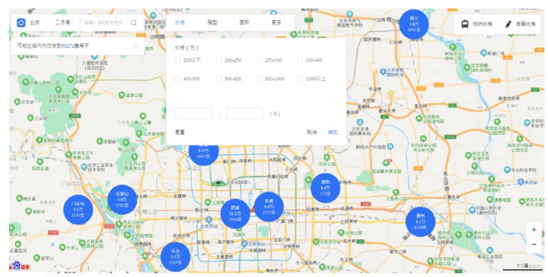

Fig. 2. Baidu Map Example

function will include searching house by map, query the history price if time permits we want to add more functions to it, and publish the app on google play. But in this project, we will first try to build a demo using the data collected from China real estate market first.

- Prospect:
  Geographic Information System can be applied to many different fields [7]. Integrating map with different data can shed light on various industries, the real estate is just an example of it. Considering Canada proposes a huge amount of forests, lakes, and other resources, We assume that GIS can be of great benefit to monitoring the resources better.

B. *WebGIS Service Analysis*

With the development and popularization of network technology, GIS based on local area network, wide area network and Internet—network GIS has subsequently become a research hotspot and an important development direction of GIS, and WebGIS is a typical Internet-based network GIS. The organic combination of GIS technology and www technology is a computer system that transmits, stores, processes, analyzes, displays and applies geospatial information under the Internet environment. The main functions of WebGIS are map functions, data acquisition, geographic data sharing, spatial analysis, etc. Compared with traditional GIS, WebGIS has the characteristics of cross-platform, simple and easy operation, easy data sharing, and wide application.

Research on WebGIS abroad is relatively early. In 1993, a university in Norway established its own map Web server [8]. In 1994, the concept of WebGIS was proposed [9]. Subsequently, a number of major GIS vendors such as ESRI, MapInfo, and Autodesk have successively researched and released their own WebGIS solutions and platform software, such as ESRI's ArcGIS, MapInfo's Map Info Proserver, and Autodesk's series of products, which are available in all sectors of society. Many applications. The GIS technology research started a bit later than other computer vision related topics, but after the continuous efforts of many geographic information scholars and technicians, GIS technology research, development and application have made great progress. In terms of WebGIS research and development, domestically, following the pace of foreign countries, many geographic information system products have soon emerged, such as SuperMap of the Chinese Academy of Sciences and GeoStar of Wuhan Gonow. However, these softwares generally have problems such as high cost and difficult maintenance.

The origin of open source GIS can be traced back to 1978 [10], when the US Department of the Interior created the MOSS system (Map Overlay and Statistics System) to track and evaluate the impact of mining development environments, wild plants, wildlife and their migration methods. Soon afterwards, GRASS (Geographical Resource Analysis Support System) was born, which is mainly used to process raster, topological vector, image and graphic resources. By 1996, GeoTools was born [11]. Four years later, the cross-platform geographic information library GDAL (Geospatial Data Abstraction Library) appeared, enabling GIS program applications to support different data formats. In 2001 [12], PostGIS was born, allowing spatial data to be stored in the Postgres database. In the same year, GeoServer was created, based on the Java language, to publish spatial data as a standard Web service. In 2002, QGIS was released [13], and it is known as the originator of open source desktop GIS. Later, with the promotion of OSGeo [14], a relatively complete production line has been formed. Under the trials of some small and medium-sized enterprises and GIS research and development enthusiasts, a small part of the commercialization is now possible.

Benefiting from the attention paid to open source in recent years, a large number of excellent open-source GIS software has emerged under the promotion of OpenGIS, including desktop GIS software uDig, QGIS, KosmoDesktop, etc.; browser and desktop clients like OpenLayers, MapBuilder; open-source database such as PostGIS, MySQL, etc.

In terms of protocols, the WebGIS Service can be categorized as follows [15]:

- Object based RPC:
  A protocol that one program can use to request a service from a program located in another computer on a network without having to understand the network's details.
- MOM based RPC:
  A type of software product that enables message distribution over complex IT systems.
- SOAP based Web Service:
  A message protocol that allows distributed elements of an application to communicate.
- Restful based Web Service:
  An architectural style for an application program interface (API) that uses HTTP requests to access and use data.

C. *Front-end Framework Analysis*

To develop a fully functional project, we decided to seperate front-end development from the back-end. To achieve this, we choose LayUI [1], which is a front-end UI framework written with its own module specifications, following the writing and organization of native HTML/CSS/JS, with extremely low thresholds, and ready to use. It is extremely simple on the outside, yet full on the inside. It is light in size and abundant in components. Every detail from the core code to the API has been meticulously crafted, making it very suitable for rapid interface development.

LayUI is more for back-end developers, so in terms of organizational form, it resolutely adopted the browser-hosted AMD-like module management method several years ago, but it is not limited to the rules and regulations of CommonJS. It has its own Mode, more lightweight and simple. Layui is defined as ”classic modularity”, not to deliberately emphasize the ”module” concept itself, but to deliberately avoid the current mainstream solutions of the JS community, and try to interpret the efficiency in the simplest possible way! Its so-called classic lies in its obsession with returning to the basics. It organizes modules in a way generally recognized by current browsers. LayUI believes that this lightweight organization can still fill many scenarios outside of WebPack. Therefore, it insists on adopting classic modularity, which is precisely what allows people to avoid the complicated configuration of tools and return to the original HTML/CSS/JavaScript itself.

IV. System Design

A. *Overall Pipeline*

As shown in Fig 3.

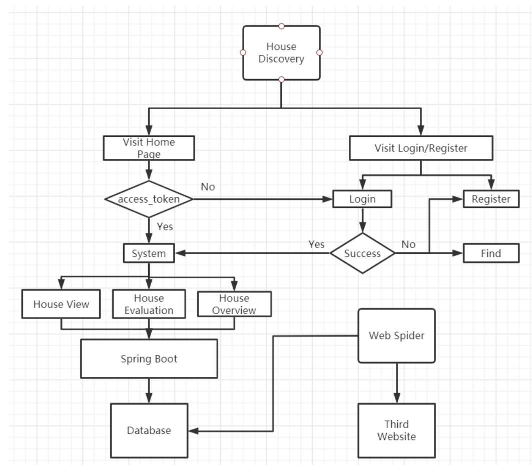

Fig. 3. Overall Pipeline

B. *Basic Client Pipeline*

The Client Front provides a cloud platform in the browser, establish communication with the background, monitor user events, generate and send corresponding requests and actions, receive the response from the server, and then render the page. As shown in Fig 7.

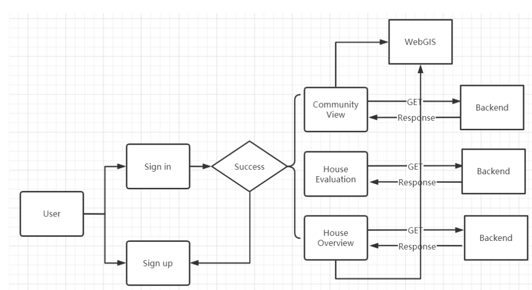

Fig. 4. Basic Client Pipeline

C. Basic Server Pipeline

The Server is designed to receive actions and requests sent by the client, process and respond, and at the same time respond to the operation and processing of the database. As shown in Fig 5.

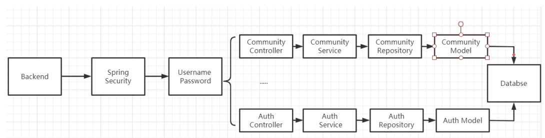

Fig. 5. Basic Server Pipeline

D. Web Spider Pipeline

The Web Spider can first analyze the web page structure, confirm the xml path of the data that needs to be scrapped, and then use the crawler framework webmagic to continuously send requests to get the response pages, extract the specified elements of the page, and then store the extracted data in the database. As shown in Fig 6.

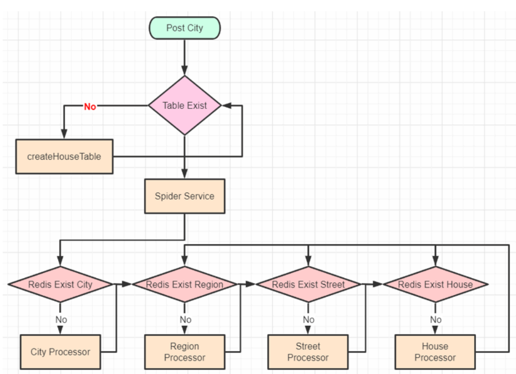

Fig. 6. Web Spider Pipeline

E. Structure Design

Users are divided into ordinary users and administrator users. Ordinary users log in to the system with the role, and then retrieve and evaluate housing information; administrator users log in to the system with the role of admin, and can update the building information in the neighborhood.

Ordinary users log in to the system with their username or email, and after passing the authentication in the background, the json web token is returned, and then the user can view the information of all the houses for sale in the community through the form, and click on a row of the form to further filter the nearby building information and display the map. , You can set the name and distance of the nearby buildings that need to be retrieved. The front-end first retrieves and draws through the map API, and then sends a request to the back-end to see if the information associated with the cell needs to be updated.

The administrator can uniformly update the building information near the cell, select the building type, send a request to the front end to obtain a list containing all cell information, and then traverse the list regularly, obtain the latitude and longitude and nearby building information through the map API, and then send the result Store and update data in the background.

F. Client Module

This Module contains user login and registration modules, user role registration, and the administrator is currently operated and set by the database administrator. As shown in Fig 7.

- Homepage: Residential housing price and geographic information page, including data form module and map module.
- House analysis: house analysis page, including data form module and heat map and radar map module drawn by echarts.
- User information: user information module
- System management module: user, role, menu management module.

G. Server Module

The server mainly includes user registration, login, spring security authentication and authorization modules, as well as modules corresponding to different models, such as model, DTO, controller, service, repository, etc. corresponding to

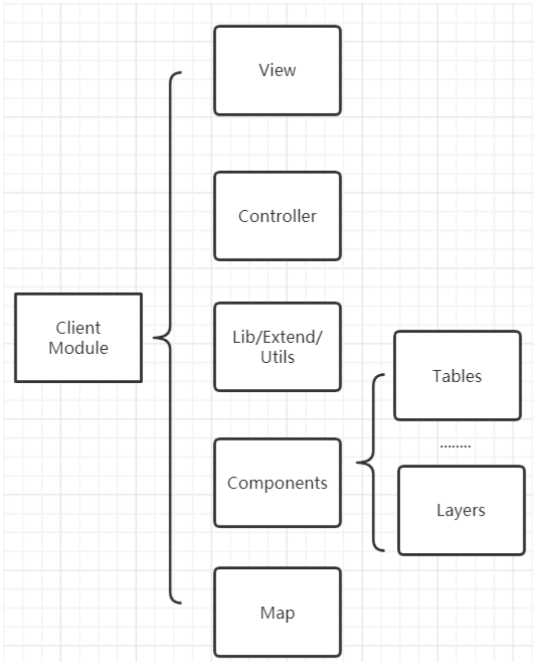

Fig. 7. Client Module

house. Different models involve different queries and database modifications. As shown in Fig 8

H. Web Spider Module

The web spider module contains different processors to extract information such as city, house, region, street, etc., as well as modules to start crawlers regularly, and various tool modules.

V. Data Structure

This section briefly demonstrates the relationships between data and how they connect with each other.

Users and Roles are a many-to-many relationship. As shown in Fig 9.

House has a many-to-many relationship with data that has information about bus, shop, school, hospital, etc. As shown in Fig 10, 12, 11 and 13.

VI. Experiment Results and Discussion

A. User Login

The login page has As shown in Fig 14, the UI was designed to fully emphasize the purpose and aesthetic aspect of our

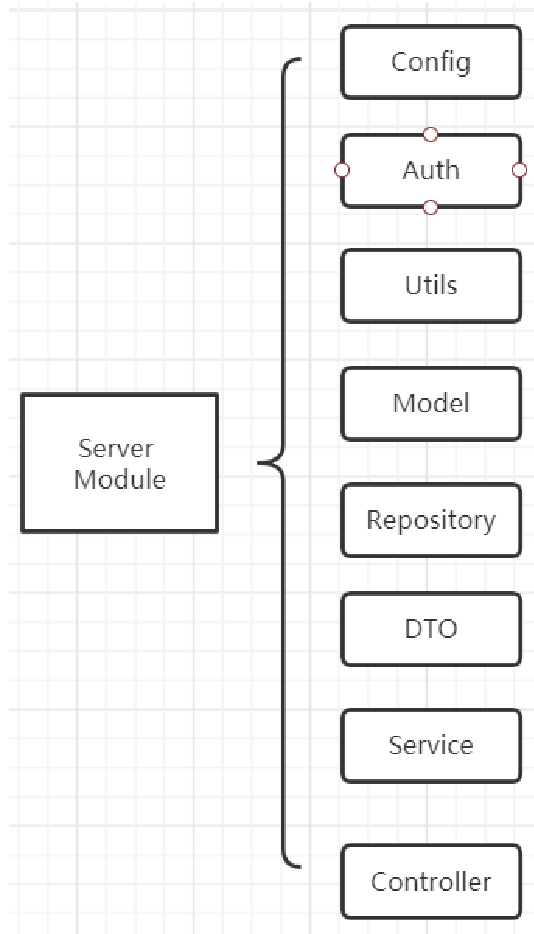

Fig. 8. Server Module

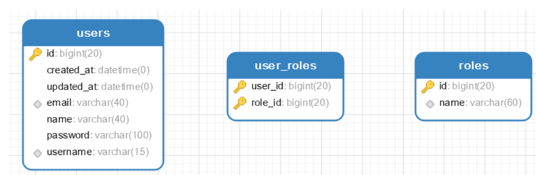

Fig. 9. User-Role Relationship

project. When a user log into the system, a message box will return a comprehensive message that clearly notice the users their login status. If the server is shutdown, a corresponding error message will return with a vibrating message box.

B. House Information Query

As shown in Fig 15. Currently the data we scrapped from Lianjia.com contains hundreds of thousands of house information, along with the bus, school, shopping mall information around each house location. Thus the application will take a

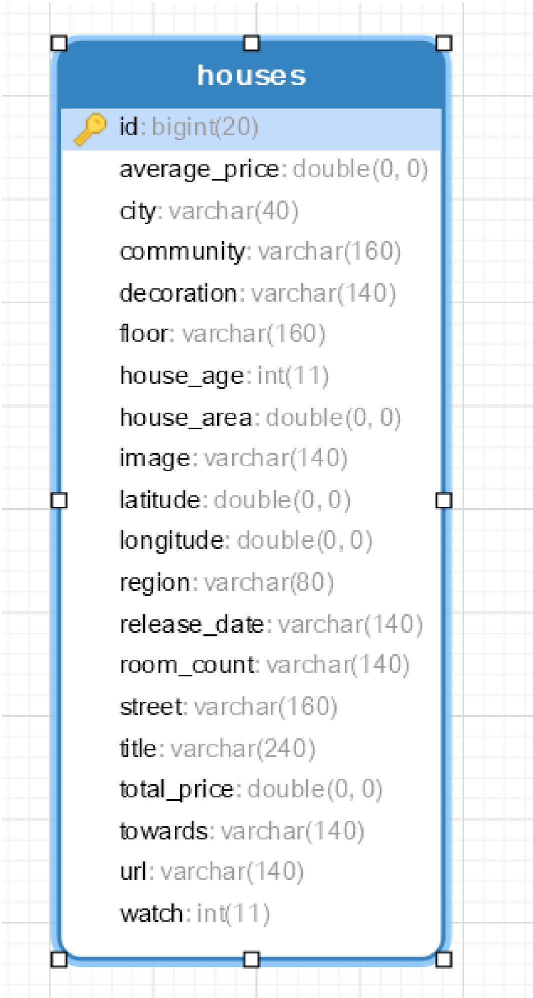

Fig. 10. House Data Structure

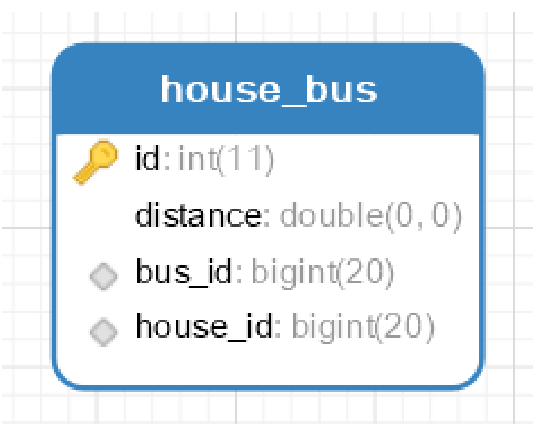

Fig. 11. House-Bus Relationship

certain amount of time to load the house information successfully. This can be improved by only query top 10 pages of the

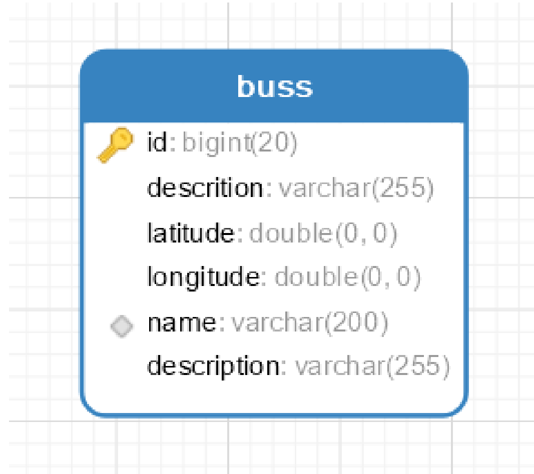

Fig. 12. Bus Data Structure

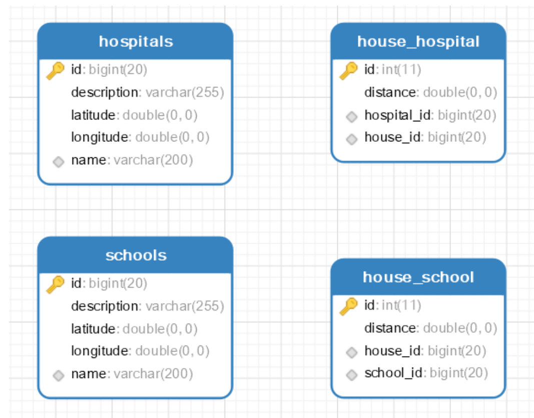

Fig. 13. Hospital-School Relationship

most relevant house information, which can obviously reduce the heavy load of the system's data processing procedure.

C. Interactive Map

As shown in Fig 16. Currently the program is able to show the amount of house information displayed at the table on the right of the interface. Further improvement can be implemented by adding surrounding bus, school etc. with different color patterns on the interactive map.

D. Heat Map

As shown in Fig 18. The heat map demonstrates the overall house price distribution depending on the color within the displayed vicinity. This function can be further improved by labeling which color represents price tag level, so that the

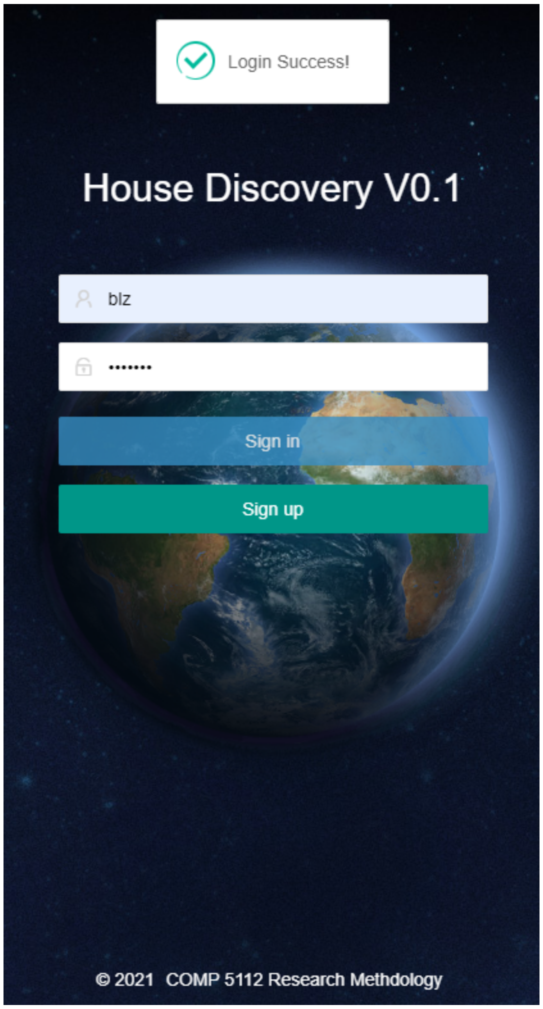

Fig. 14. House Discovery login

Community List

city	regio...	street	community	average_tot...	average_pri...
上海	闵行	七宝	七宝老街	556.5	106375.5
上海	闵行	七宝	万兆家园(一期)	608.285714...	79552.2857...
上海	闵行	七宝	万兆家园(五期)	654.5	76255.5
上海	闵行	七宝	万兆家园(四期)	660	75628
上海	闵行	七宝	万泰公寓	696	63708.3333...
上海	闵行	七宝	万泰花园	603.454545...	89057.6363...
上海	闵行	七宝	万科七宝国际	256.916666...	58054
上海	闵行	七宝	万科优诗美地	975	77721.3333...
上海	闵行	七宝	万科凭栏苑	1170	84488.6666...
上海	闵行	七宝	万科城市花园	660.041666...	69694.0833...
上海	闵行	七宝	万科城花新园	1028	89067.6363...
上海	闵行	七宝	万科朗润园	928.636363...	85184.4545...
上海	闵行	七宝	万科桂馨苑	1188	82666.5

1 2 3 ... 776 到第 1 页 确定 共 7752 条 10 条/页

Fig. 15. House Information Query

Fig. 16. Interactive Map

users can get a more comprehensive understanding on what price level the heat map is displaying.

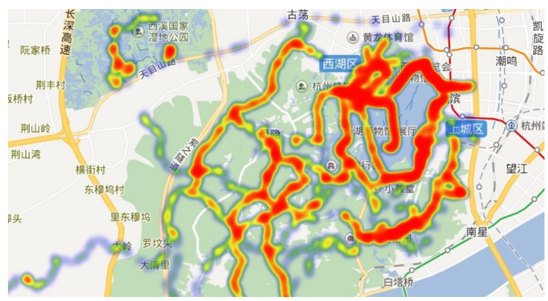

Fig. 17. Heat Map

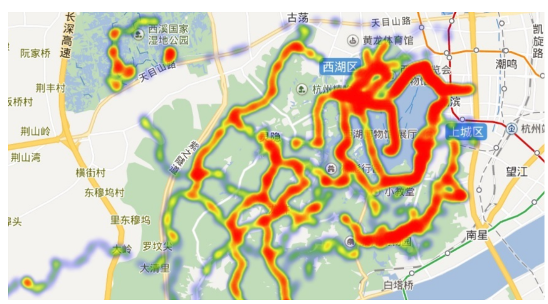

Fig. 18. Heat Map

VII. Conclusions and Future Works

For this project, we developed a program based on WebGIS for House Discovery. Completed the system, database, interface, and interface design.

We have successfully implemented basic functions such as user login and registration. At the same time, we use Spring Security's components to implement user authentication system to ensure the user's login authority and security. The user authentication system can also be used to distinguish between ordinary users and administrator roles. The method of defining user roles is safer than the common database field roles.

This project also realizes the function of real-time acquisition of house data. Using Java Springboot components, we have developed a Web Spider that can update house information online every day, so that if the house is sold or the information is no longer public, the user can use our system that obtains up-to-date and useful housing information.

This program can display the specific location of the house on the map according to the acquired house information. This function has also added a number of extensions, such as using bus, school and other surrounding landmark building information to obtain the required house information.

At present, due to time constraints and development feasibility considerations during the project analysis phase, we switched from the python Flask development to Java Springboot and Spring Security in the background development process to implement user authentication. In addition, because the Hibernate component is more suitable for single-table operation , It will encounter many tricky problems when dealing with multiple tables and complex database queries, and it encounters a bottleneck in the process of program implementation. We spent a lot of time debugging and finding solutions. Therefore, this project still has a lot of room for improvement. For example, Spring Security is more suitable for the development of Role-based access system with front-end, admin roles can be used for different users through the system management page The role and permission configuration.

In the future we will continue to polish the existing modules and to develop unfinished functions as we explained above in detail. This project can expand to a fully commercial-capable model that helps people to find their desired house.

References

[1] Introduction to layui Part 1 (Le byte architecture, big data). [Online]. Available: https://developpaper.com/introduction-to-layui-part-1-le-byte-architecture-big-data/. [Accessed: April.3, 2020].

[2] GEORG KOSTERS, BERND-UWE PAGEL HANS-WERNER SIX (1997) GIS-application development with GEOOOA, International Journal of Geographical Information Science, 11:4, 307-335, DOI: 10.1080/136588197242293

[3] Congalton, Russell G. "Exploring and evaluating the consequences of vector-to-raster and raster-to-vector conversion." Photogrammetric Engineering and Remote Sensing 63.4 (1997): 425-434.

[4] Mann, Ronald J., and Seth R. Belzley. "The promise of internet intermediary liability." Wm. Mary L. Rev. 47 (2005): 239.

[5] Baidu Maps [Online]. Available: https://map.baidu.com/. [Accessed: April.3, 2020].

[6] Trulia: Real Estate Listings, Homes For Sale, Housing Data [Online]. Available: https://www.trulia.com/. [Accessed: March.23, 2020].

[7] Chang, Kang-Tsung. "Geographic information system." International Encyclopedia of Geography: People, the Earth, Environment and Technology (2016): 1-10.

[8] Peterson, Michael P. "Trends in Internet map use." Proceedings of the 19th ICA Conference, Ottawa. Vol. 1. No. 10. 1997.

[9] Agrawal, Sonam, and R. D. Gupta. "Web GIS and its architecture: a review." Arabian Journal of Geosciences 10.23 (2017): 1-13.

[10] Bourgeois, Pierre, et al. "GEOGRAPHIC INFORMATION SYSTEM APPLICATIONS FOR MARSH MANAGEMENT PLANS." Marsh Management in Coastal Louisiana: Effects and Issues: Proceedings of a Symposium. US Department of the Interior, Fish and Wildlife Service, Research and Development, 1989.

[11] Turton, Ian. "Geo tools." Open source approaches in spatial data handling. Springer, Berlin, Heidelberg, 2008. 153-169.

[12] Vasin, Yu G., and Yu V. Yasakov. "Object-oriented topological management system of spatially-distributed databases." Pattern Recognition and Image Analysis 26.4 (2016): 734-741.

[13] Kurt Menke, G. I. S. P., et al. Mastering QGIS. Packt Publishing Ltd, 2016.

[14] Hall, George Brent. Open source approaches in spatial data handling. Ed. Michael G. Leahy. Vol. 2. Berlin: Springer, 2008.

[15] Kolev, Alexander. "AstroWeb Astroinformatics Project and Comparison of the WEB-GIS Protocol Standards." (2015).